

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-101124

(P2002-101124A)

(43) 公開日 平成14年4月5日 (2002.4.5)

(51) Int.Cl.⁷

識別記号

F I

テ-マコ-ド* (参考)

H 0 4 L 12/56

H 0 4 L 11/20

1 0 2 A 5 K 0 3 0

12/28

11/00

3 1 0 D 5 K 0 3 3

12/66

11/20

B 5 K 0 3 4

29/08

13/00

3 0 7 C

審査請求 未請求 請求項の数11 O L (全 14 頁)

(21) 出願番号

特願2000-288873(P2000-288873)

(22) 出願日

平成12年9月22日 (2000.9.22)

(71) 出願人 000003078

株式会社東芝

東京都港区芝浦一丁目1番1号

(72) 発明者 室野 隆博

東京都府中市東芝町1番地 株式会社東芝
府中事業所内

(74) 代理人 100058479

弁理士 鈴江 武彦 (外6名)

Fターム(参考) 5K030 HD03 KA03 KX12 KX13 KX18

LC08 LC11 MB15

5K033 CB06 CB08 DA01 DB13 DB18

5K034 FF11 HH42 HH50 HH56 HH64

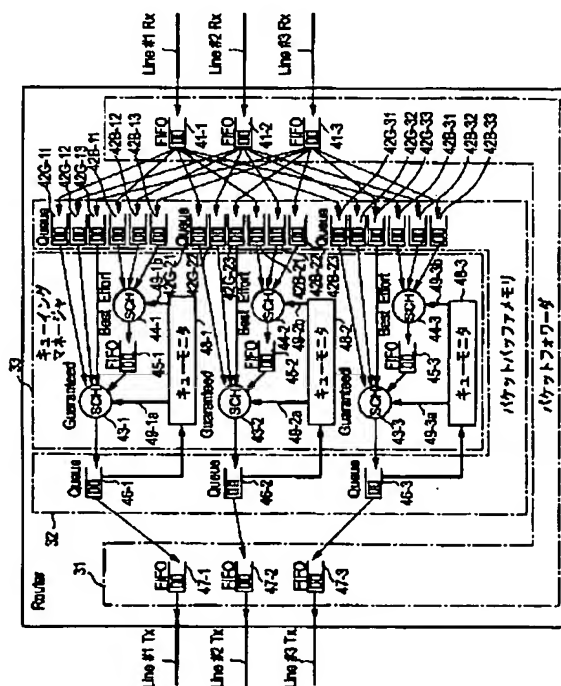
MM08

(54) 【発明の名称】 ネットワーク間中継装置

(57) 【要約】

【課題】 可変長パケットを中継するネットワーク間中継装置において、重要度の高いデータの最小帯域を保証する。

【解決手段】 本発明は、入力インターフェイス毎に設けられた保証トラフィック用スケジューラ(43-1~43-3)及びベストエフォートトラフィック用スケジューラ(44-1~44-3)がそれぞれ保証トラフィック用入力キュー(42G-11~42G-13、42G-21~23、42G-31~33)及びベストエフォートトラフィック用入力キュー(42B-11~13、42B-21~23、42B-31~33)に格納されたキューの取り出しレートを制御する。これにより、インターフェイス毎に保証帯域が保証され、かつ保証帯域内において各入力インターフェイスからの保証データの取り出しレートを設定することができる。



【特許請求の範囲】

【請求項1】 入力インターフェイス毎に設けられ、前記入力インターフェイスに対応する出力インターフェイスに出力される所定の品質が要求される保証データを格納する保証トラフィック用入力キューと、前記入力インターフェイス毎に設けられ、入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納するベストエフォートトラフィック用入力キューと、前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートをそれぞれ制御するベストエフォートトラフィック用スケジューラと、前記ベストエフォートトラフィック用スケジューラによって前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートをそれぞれ制御する保証トラフィック用スケジューラと、前記出力インターフェイス毎に設けられ、前記保証トラフィック用スケジューラによって取り出されたデータをそれぞれ格納する出力キューとを具備することを特徴とするネットワーク間中継装置。

【請求項2】 前記出力インターフェイス毎に設けられ、前記出力キューに格納されたデータに基づいて、対応する出力インターフェイスの輻輳を検出するキューモニターをさらに具備し、前記ベストエフォートトラフィック用スケジューラ及び前記保証トラフィック用スケジューラは、対応する前記キューモニターが輻輳を検出した場合に、取り出しレートを変更することを特徴とする請求項1記載のネットワーク間中継装置。

【請求項3】 各入力インターフェイスから入力されるパケットデータを出力インターフェイス及び通信品質クラスに基づいて分類し、この分類に基づいて前記パケットデータを前記入力インターフェイス毎に設けられた前記保証トラフィック用入力キュー及び前記ベストエフォートトラフィック用入力キューに出力するパケットフォワーダをさらに具備することを特徴とする請求項1又は請求項2記載のネットワーク間中継装置。

【請求項4】 前記キューモニターによって対応する出力インターフェイスの輻輳が検出された場合における前記保証トラフィック用スケジューラの前記保証トラフィック用入力キューに格納された保証データの取り出しレートは、下記式に基づいて決定されることを特徴とする請求項2記載のネットワーク間中継装置。

(対応する入力インターフェイスのトラフィックが占める保証帯域) / (対応する出力インターフェイスの採りうる最大帯域)

【請求項5】 前記キューモニターによって対応する出

力インターフェイスの輻輳が検出された場合における前記保証トラフィック用スケジューラの前記ベストエフォートトラフィック用入力キューに格納された保証データ以外のデータの取り出しレートは、下記式に基づいて決定されることを特徴とする請求項2記載のネットワーク間中継装置。

$$\{ (\text{対応する出力インターフェイスのとり得る最大帯域}) - (\text{対応する出力インターフェイスの保証帯域}) \} / (\text{対応する出力インターフェイスのとり得る最大帯域})$$

【請求項6】 前記キューモニターによって対応する出力インターフェイスの輻輳が検出された場合におけるベストエフォートトラフィック用スケジューラの前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートは、下記式に基づいて決定されることを特徴とする請求項2記載のネットワーク間中継装置。

$$\{ (\text{対応する入力インターフェイスの採りうる最大帯域}) - (\text{対応する入力インターフェイスの保証帯域}) \} / \Sigma \{ (\text{入力インターフェイスの採りうる最大帯域}) - (\text{入力インターフェイスの保証帯域}) \}$$

【請求項7】 入力インターフェイスに対応する出力インターフェイスに出力される保証データを格納させる第1の格納手段と、

前記入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納させる第2の格納手段と、

前記第2の格納手段によって格納された前記保証データ以外のデータの取り出しレートを個別に制御するベストエフォートトラフィック用スケジューラと、

前記ベストエフォートトラフィック用スケジューラによって取り出された前記保証データ以外のデータと、前記第1の格納手段によって格納された保証データとの取り出しレートをそれぞれ個別に制御する保証トラフィック用スケジューラと、

前記保証トラフィック用スケジューラによって取り出されたデータを前記入力インターフェイスに対応する出力インターフェイスに出力する手段とを具備することを特徴とするコンピュータ読み取り可能な情報記録媒体。

【請求項8】 入力インターフェイス毎に設けられ、前記入力インターフェイスに対応する出力インターフェイスに出力される所定の品質が要求される保証データを格納する保証トラフィック用入力キューと、前記入力インターフェイス毎に設けられ、入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納するベストエフォートトラフィック用入力キューと、前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートを個別に制御するベストエフォートトラフィック用スケジ

ューラと、
 前記ベストエフォートトラフィック用スケジューラによって前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートをそれぞれ個別に制御する保証トラフィック用スケジューラと、
 前記出力インターフェイス毎に設けられ、前記保証トラフィック用スケジューラによって取り出されたデータをそれぞれ格納する出力キューと、
 前記出力インターフェイス毎に設けられ、前記出力キューに格納されたデータに基づいて、対応する出力インターフェイスの輻輳を検出するキューモニターとを具備し、
 前記キューモニターが対応する出力インターフェイスの輻輳を検出した場合に、対応する前記保証トラフィック用スケジューラ及び前記ベストエフォートトラフィック用スケジューラに輻輳を通知し、
 前記対応するキューモニターから輻輳を通知された場合に、前記ベストエフォートトラフィック用スケジューラは前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートを変更し、前記保証トラフィック用スケジューラは、前記ベストエフォートトラフィック用スケジューラによって前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートを変更することを特徴とするネットワーク間中継方法。
 【請求項9】 入力インターフェイス毎に設けられ、前記入力インターフェイスに対応する出力インターフェイスに出力される所定の品質が要求される保証データを格納する保証トラフィック用入力キューと、
 前記入力インターフェイス毎に設けられ、入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納するベストエフォートトラフィック用入力キューと、
 前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートをそれぞれ制御する保証トラフィック用スケジューラと、
 前記出力インターフェイス毎に設けられ、前記保証トラフィック用スケジューラによって取り出されたデータをそれぞれ格納する出力キューとを具備することを特徴とするネットワーク間中継装置。
 【請求項10】 前記ベストエフォートトラフィック用スケジューラは、前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータを

パッケージ単位で取り出すことを特徴とする請求項1記載のネットワーク間中継装置。

【請求項11】 前記保証トラフィック用スケジューラは、前記保証トラフィック用入力キューに格納された前記保証データをパッケージ単位で取り出すことを特徴とする請求項9記載のネットワーク間中継装置。

【発明の詳細な説明】

【0001】

10 【発明の属する技術分野】本発明は、複数のLAN（ローカルエリアネットワーク）およびWAN（広域ネットワーク）同士を相互に接続し、それらの間でデータを中継するネットワーク間中継装置に関するものである。

【0002】

【従来の技術】図5は、ネットワーク間中継装置であるルータの一般的な内部ブロック図である。

【0003】同図に示すように、ネットワーク間中継装置は、1ポート以上のWANまたはLANのラインインタフェース（Line #1～Line #3）を持っており、各ポートは、PMD（Physical Layer Medium Dependent）およびPHY（Physical Layer）11-1～11-3の各機能により構成される。

【0004】また、ラインインタフェース毎にデータリンク層に相当するデータリンクコントローラ12-1～12-3が設けられており、ラインとのパケット送受信処理が行われる。図5においては、便宜的にラインインタフェースを3ポート設けている。

【0005】パケットフォワード13は、データリンクコントローラ12-1～12-3からパケットを受信すると、パケットのアドレス、例えばIP（Internet Protocol）アドレスに基づいて、転送すべきラインインタフェースを判断し、該当する送信ラインインタフェース宛のパケットとして中継処理し、キューイングマネージャ15経由でパケットバッファメモリ14に書き込む。

【0006】また、逆にデータリンクコントローラ12-1～12-3に送信する場合も、キューイングマネージャ15経由でパケットバッファメモリ14から読み出されたパケットについて、同様に中継処理を行いデータリンクコントローラ12-1～12-3に対して送信要求を発行する。

【0007】上記中継処理には、IPルータが通常行うようなパケットヘッダのフィールドのチェックや追加／削除／書き換え等の処理が含まれる。但し、ラインインタフェースから受信した時とラインインタフェースへ送信する時とでは処理内容が異なる場合がある。

【0008】パケットバッファメモリ14は、パケットフォワード13からキューイングマネージャ15経由で受信したパケットを一時的に保持するためのメモリであ

る。

【0009】キューイングマネージャ15は、中継されるパケットのキューを管理する機能を持ち、パケットバッファメモリ14に格納されているパケットデータをパケットフォワード13に対して送信する順序およびタイミングをパケット毎に調整する。

【0010】また、キューイングマネージャ15はパケットフォワード13が送受信パケットデータをパケットバッファメモリ14との間でリード/ライトするためのDMAコントローラ機能を持ち、キューイングマネージャ15内のスケジューラとこのDMAコントローラがパケットバッファメモリをアクセスする際のアクセス権の調停を行うアービタ機能を持つ。

【0011】図6は、従来のネットワーク間中継装置のパケット処理系を示す図である。なお、図5と同一部分には、同一符号を付して説明する。

【0012】また、同図において、ラインインタフェースのデータリンクコントローラから外側の部分については省略しており、便宜上ラインインタフェースの各ポートを送信ラインと受信ラインに明示的に分けて示している。

【0013】同図に示すように、入力インターフェイス(Line#1Rx~Line#3Rx)毎にFIFOバッファ(FIFO-1~FIFO-3)が設けられている。このFIFOバッファ(FIFO-1~FIFO-3)には、対応する入力インターフェイスから入力されるパケットデータが順次格納される。

【0014】パケットフォワード13は、FIFOバッファ(FIFO-1~FIFO-3)に格納されたパケットデータを解析して、パケットデータを、その宛先の出力インターフェイス(Line#1Tx~Line#3Tx)毎に設けられたキュー(3-11~3-13、3-21~3-23、3-31~3-33)にキューイングする。

【0015】なお、上記キュー(3-11~3-13、3-21~3-23、3-31~3-33)は、パケットバッファメモリ14に設けられている。

【0016】例えば、パケットフォワード13は、入力インターフェイス(Line#1Rx)から入力され、FIFO-1に格納されたパケットデータの宛先が出力インターフェイス(Line#1Tx)である場合には、そのパケットデータをキュー3-11にキューイングする。

【0017】また、入力インターフェイス(Line#1Rx)から入力され、FIFO-1に格納されたパケットデータの宛先が出力インターフェイス(Line#2Tx)である場合には、そのパケットデータをキュー3-21にキューイングする。

【0018】出力インターフェイス毎に設けられたキュー(3-11~3-13、3-21~3-23、3-3

1~3-33)にキューイングされたパケットデータは、対応するスケジューラ4-1~4-3によってパケットが取り出され、パケットバッファメモリ14に出力インターフェイス(Line#1Tx~Line#3Tx)毎に設けられたキュー5-1~5-3にキューイングされる。

【0019】図7は、スケジューラ4-1~4-3の動作を説明するための状態遷移図である。

【0020】ここでは、入力インターフェイスLine#1Rxに対応して設けられたスケジューラ4-1を例にとり説明する。

【0021】まず、スケジューラ4-1は、入力インターフェイスLine#1Rxに対応するキュー3-1に格納されたパケットデータの取り出しを行なう(S1)。

【0022】そして、キュー3-1に格納された先頭のパケットデータを全て取り出し、すなわち、キュー3-1の先頭パケットデータを最後まで取り出し終わった場合に、S2の処理に移行する。

【0023】S2においては、スケジューラ4-1は、入力インターフェイスLine#2Rxに対応するキュー3-2に格納されたパケットデータの取り出しを行なう(S1)。

【0024】そして、S1と同様に、キュー3-2の先頭パケットデータを最後まで取り出し終わった場合に、S3の処理に移行する。

【0025】S3においても同様に、スケジューラ4-1が、キュー3-13に格納されたパケットデータの取り出しを行ない、キュー3-13の先頭パケットデータを最後まで取り出し終わった場合に、S1の処理に戻る。

【0026】なお、ここでは、スケジューラ4-1について説明したが、スケジューラ4-2、4-3についても、対応するキューについて同様の処理が行なわれる。また、各状態で該当するキューにパケットデータが存在しない場合は、その状態での処理をスキップし、次の状態に遷移する。

【0027】

【発明が解決しようとする課題】しかしながら、上述の従来のネットワーク間中継装置には以下のような問題があった。すなわち、ネットワーク間において伝送されるデータには、重要度が高いデータ及び低いデータがあるが、従来のネットワーク間中継装置においては、これらデータの重要度(要求される通信品質)が考慮されていない。

【0028】したがって、ネットワークが輻輳状態になった場合に、重要度の低いデータ(要求通信品質の低いデータ)のみならず重要度の高いデータ(要求通信品質の高いデータ)の中継も困難になってしまうという問題があった。

【0029】なお、ネットワーク間中継装置の1つである可変長パケットスイッチとしては、IEEE802.1P準拠のイーサネット（登録商標）スイッチのようにパケットに付けられた優先度に基づいて優先度の高いパケットからネットワーク上に送出して特定トラフィックの通信品質を高める優先制御機能を持つものがある。

【0030】しかしながら、このような可変長パケットスイッチもパケットの送出の順番を優先度に基づいて変更しているのみで、重要度の高いデータ（要求通信品質の高いデータ）の通信品質を保証するものではない。

【0031】また、ネットワーク間中継装置の1つである固定長セルを交換するATM Cellスイッチでは、最小帯域を保証するGFR（Guaranteed Frame Rate）QOSクラスを実現しているものがある。

【0032】しかしながら、このATM Cellスイッチは、固定長セルに関するものであり、可変長パケットを中継するネットワーク間中継装置に適用することはできない。

【0033】本発明は、上記実情に鑑みてなされたものであり、可変長パケットを中継するネットワーク間中継装置において、重要度の高いデータの最小帯域を保証することができるネットワーク間中継装置を提供することを目的とする。

【0034】

【課題を解決するための手段】したがって、まず、上記目的を達成するために、本発明の第1の発明は、入力インターフェイス毎に設けられ、前記入力インターフェイスに対応する出力インターフェイスに出力される所定の品質が要求される保証データを格納する保証トラフィック用入力キューと、前記入力インターフェイス毎に設けられ、入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納するベストエフォートトラフィック用入力キューと、前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートを制御するベストエフォートトラフィック用スケジューラと、前記ベストエフォートトラフィック用スケジューラによって前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートをそれぞれ制御する保証トラフィック用スケジューラと、前記出力インターフェイス毎に設けられ、前記保証トラフィック用スケジューラによって取り出されたデータをそれぞれ格納する出力キューとを具備することを特徴とするネットワーク間中継装置である。

【0035】このような発明によれば、保証トラフィック用スケジューラ及びベストエフォートトラフィック用スケジューラがそれぞれ保証トラフィック用入力キュー

及びベストエフォートトラフィック用入力キューに格納されたキューの取り出しレートを制御することができるので、インターフェイス毎に保証帯域が保証され、かつ保証帯域内において各入力インターフェイスからの保証データの取り出しレートを設定することができる。

【0036】また、本発明の第2の発明は、第1の発明において、前記出力インターフェイス毎に設けられ、前記出力キューに格納されたデータに基づいて、対応する出力インターフェイスの輻輳を検出するキューモニターをさらに具備し、前記ベストエフォートトラフィック用スケジューラ及び前記保証トラフィック用スケジューラは、対応する前記キューモニターが輻輳を検出した場合に、取り出しレートを変更することを特徴とするネットワーク間中継装置である。

【0037】このような発明によれば、キューモニターによって出力インターフェイスの輻輳が検出された場合に、保証トラフィック用スケジューラ及びベストエフォートトラフィック用スケジューラがそれぞれ保証トラフィック用入力キュー及びベストエフォートトラフィック用入力キューに格納されたキューの取り出しレートを制御する。これにより、ネットワークの輻輳が検出された場合に、保証データの通信品質を確保することができる。

【0038】さらに、本発明の第3の発明は、第1の発明又は第2の発明において、各入力インターフェイスから入力されるパケットデータを出力インターフェイス及び通信品質クラスに基づいて分類し、この分類に基づいて前記パケットデータを前記入力インターフェイス毎に設けられた前記保証トラフィック用入力キュー及び前記ベストエフォートトラフィック用入力キューに出力するパケットフォワードをさらに具備することを特徴とするネットワーク間中継装置である。

【0039】さらに、本発明の第4の発明は、第2の発明において、前記キューモニターによって対応する出力インターフェイスの輻輳が検出された場合における前記保証トラフィック用スケジューラの前記保証トラフィック用入力キューに格納された保証データの取り出しレートは、下記式に基づいて決定されることを特徴とする請求項2記載のネットワーク間中継装置ネットワーク間中継装置である。

【0040】（対応する入力インターフェイスのトラフィックが占める保証帯域）／（対応する出力インターフェイスの採りうる最大帯域）

さらに、本発明の第5の発明は、第2の発明において、前記キューモニターによって対応する出力インターフェイスの輻輳が検出された場合における前記保証トラフィック用スケジューラの前記ベストエフォートトラフィック用入力キューに格納された保証データ以外のデータの取り出しレートは、下記式に基づいて決定されることを特徴とするネットワーク間中継装置である。

【0041】{(対応する出力インターフェイスのとり得る最大帯域) - (対応する出力インターフェイスの保証帯域)} / (対応する出力インターフェイスのとり得る最大帯域)

さらに、本発明の第6の発明は、前記キューモニターによって対応する出力インターフェイスの輻輳を検出された場合におけるベストエフォートトラフィック用スケジューラの前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートは、下記式に基づいて決定されることを特徴とするネットワーク間中継装置である。

【0042】{(対応する入力インターフェイスの採りうる最大帯域) - (対応する入力インターフェイスの保証帯域)} / \sum {(入力インターフェイスの採りうる最大帯域) - (入力インターフェイスの保証帯域)}

さらに、本発明の第7の発明は、入力インターフェイスに対応する出力インターフェイスに出力される保証データを格納させる第1の格納手段と、前記入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納させる第2の格納手段と、前記第2の格納手段によって格納された前記保証データ以外のデータの取り出しレートを個別に制御するベストエフォートトラフィック用スケジューラと、前記ベストエフォートトラフィック用スケジューラによって取り出された前記保証データ以外のデータと、前記第1の格納手段によって格納された保証データとの取り出しレートをそれぞれ個別に制御する保証トラフィック用スケジューラと、前記保証トラフィック用スケジューラによって取り出されたデータを前記入力インターフェイスに対応する出力インターフェイスに出力する手段とを具備することを特徴とするコンピュータ読み取り可能な情報記録媒体である。

【0043】さらに、本発明の第8の発明は、入力インターフェイス毎に設けられ、前記入力インターフェイスに対応する出力インターフェイスに出力される所定の品質が要求される保証データを格納する保証トラフィック用入力キューと、前記入力インターフェイス毎に設けられ、入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納するベストエフォートトラフィック用入力キューと、前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートを個別に制御するベストエフォートトラフィック用スケジューラと、前記ベストエフォートトラフィック用スケジューラによって前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートをそれぞれ個別に制御する保証トラフィック用スケジューラと、前記出力インターフェイス毎に設けられ、前記保証トラフィック用スケジュー

ラによって取り出されたデータをそれぞれ格納する出力キューと、前記出力インターフェイス毎に設けられ、前記出力キューに格納されたデータに基づいて、対応する出力インターフェイスの輻輳を検出するキューモニターとを具備し、前記キューモニターが対応する出力インターフェイスの輻輳を検出した場合に、対応する前記保証トラフィック用スケジューラ及び前記ベストエフォートトラフィック用スケジューラに輻輳を通知し、前記対応するキューモニターから輻輳を通知された場合に、前記ベストエフォートトラフィック用スケジューラは前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータの取り出しレートを変更し、前記保証トラフィック用スケジューラは、前記ベストエフォートトラフィック用スケジューラによって前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートを変更することを特徴とするネットワーク間中継方法である。

【0044】さらに、本発明の第9の発明は、入力インターフェイス毎に設けられ、前記入力インターフェイスに対応する出力インターフェイスに出力される所定の品質が要求される保証データを格納する保証トラフィック用入力キューと、前記入力インターフェイス毎に設けられ、入力インターフェイスに対応する出力インターフェイスに出力される前記保証データ以外のデータを格納するベストエフォートトラフィック用入力キューと、前記ベストエフォートトラフィック用入力キューから取り出された前記保証データ以外のデータと、前記保証トラフィック用入力キューに格納された保証データとの取り出しレートをそれぞれ制御する保証トラフィック用スケジューラと、前記出力インターフェイス毎に設けられ、前記保証トラフィック用スケジューラによって取り出されたデータをそれぞれ格納する出力キューとを具備することを特徴とするネットワーク間中継装置である。

【0045】このような発明によれば、保証トラフィック用スケジューラ保証トラフィック用入力キューに格納されたキューの取り出しレートを制御することができるので、インターフェイス毎に保証帯域が保証され、かつ保証帯域内において各入力インターフェイスからの保証データの取り出しレートを設定することができる。

【0046】さらに、本発明の第10の発明は、第1の発明において、前記ベストエフォートトラフィック用スケジューラは、前記ベストエフォートトラフィック用入力キューに格納された前記保証データ以外のデータをバケット単位で取り出し、前記保証トラフィック用スケジューラは、前記保証トラフィック用入力キューに格納された前記保証データをバケット単位で取り出すことを特徴とするネットワーク間中継装置である。

【0047】このような発明によれば、ベストエフォ

トトラフィック用スケジューラ及び保証トラフィック用スケジューラがパケット単位でデータを読み出すので、読み出しの際に転送データが破壊されるのを防止することができる。

【0048】さらに、本発明の第11の発明は、第9の発明において、前記保証トラフィック用スケジューラは、前記保証トラフィック用入力キューに格納された前記保証データをパケット単位で取り出すことを特徴とするネットワーク間中継装置である。

【0049】このような発明によれば、保証トラフィック用スケジューラがパケット単位でデータを読み出すので、読み出しの際に転送データが破壊されるのを防止することができる。

【0050】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態に係るネットワーク間中継装置について説明する。

【0051】ネットワーク間中継装置の内部ブロック図については、図5に示したものと同様であるのでここでは、省略する。ただし、本実施の形態においては、キューイングマネージャ、パケットフォワーダなどの機能が異なる。

【0052】図1は、本発明の実施の形態に係るネットワーク間中継装置のパケット処理系を示す図である。同図において、ラインインタフェースのデータリンクコントローラから外側の部分については省略しており、また便宜上ラインインタフェースの各ポートを送信ラインと受信ラインに明示的に分けて示している。

【0053】同図に示すように、入力インターフェイス(Line#1Rx~Line#3Rx)毎にFIFOバッファ41-1~41-3が設けられている。このFIFOバッファ41-1~41-3には、対応する入力インターフェイスから入力されるパケットデータが順次格納される。

【0054】パケットフォワーダ31は、FIFOバッファ41-1~41-3に格納されたパケットデータを解析して、パケットデータを、その宛先の出力インターフェイス(Line#1Tx~Line#3Tx)毎に設けられたキュー(42G-11~42G-13、42B-11~42B-13、42G-21~23、42B-21~42B-23、42G-31~33、42B-31~33)にキューイングする。

【0055】なお、上記キュー(42G-11~42G-13、42B-11~42B-13、42G-21~23、42B-21~42B-23、42G-31~33、42B-31~33)は、パケットバッファメモリ32に設けられている。

【0056】また、上記キューは、帯域保証クラス用のキュー(42G-11~42G-13、42G-21~23、42G-31~33)及びベストエフォートクラ

ス用のキュー(42B-11~13、42B-21~23、42B-31~33)にグループ分けされている。

【0057】例えば、パケットフォワーダ31は、入力インターフェイス(Line#1Rx)から入力され、FIFO41-1に格納されたパケットデータの宛先が出力インターフェイス(Line#1Tx)である場合であって、かつ帯域保証クラス用のパケットである場合には、そのパケットデータをキュー42G-11にキューイングする。

【0058】また、入力インターフェイス(Line#1Rx)から入力され、FIFO41-1に格納されたパケットデータの宛先が出力インターフェイス(Line#2Tx)である場合であって、かつベストエフォートクラス用のパケットである場合には、そのパケットデータをキュー42B-21にキューイングする。

【0059】ベストエフォートクラス用のスケジューラ44-1~44-3は、それぞれ出力インターフェイス(Line#1Tx~Line#3Tx)毎に設けられ、対応するキュー(42B-11~42B-13、42B-21~23、42B-31~33)からそれぞれパケットデータを取り出す。

【0060】また、対応するキューモニタ48-1~48-3から出力インターフェイスの輻輳が通知された場合に、対応するキュー(42B-11~42B-13、42B-21~23、42B-31~33)からのパケットデータの取り出しレートを変更する機能を有する。なお、各キューへのパケットデータのキューイングの制御はキューイングマネージャ33が行なう。

【0061】帯域保証クラス用のスケジューラ43-1~43-3は、それぞれ出力インターフェイス(Line#1Tx~Line#3Tx)毎に設けられ、対応するキュー(42G-11~13、42G-21~23、42G-31~33)及びベストエフォートクラス用のスケジューラ44-1~44-3から送出されるパケットデータを格納するFIFOメモリ45-1~45-3からそれぞれパケットデータを取り出す。

【0062】これら帯域保証クラス用のスケジューラ43-1~43-3によって取り出されたパケットデータは、パケットバッファメモリ32に出力インターフェイス(Line#1Tx~Line#3Tx)毎に設けられたキュー46-1~46-3にキューイングされる。

【0063】FIFO47-1~47-3は、出力インターフェイス(Line#1Tx~Line#3Tx)毎に設けられ、対応するキュー46-1~46-3から送出されるパケットデータをバッファリングし、順次対応する出力インターフェイス(Line#1Tx~Line#3Tx)に出力する。

【0064】次に、本発明の実施の形態に係るネットワーク間中継装置の動作について説明する。

【0065】入力インターフェイス(Line#1Rx

～Line#3Rx)を介して受信され、パケットフォワード31により中継処理が行なわれたパケットは、パケットフォワード31内部のFIFO41-1～41-3経由で、出力インターフェイス(Line#1Tx～Line#3Tx)及び通信品質クラス(帯域保証クラス、ベストエフォートクラス)毎に振り分けられて、対応するキュー(42G-11～42G-13、42B-11～42B-13、42G-21～23、42B-21～42B-23、42G-31～33、42B-31～33)にキューイングされる。

【0066】ベストエフォートクラス用のスケジューラ(44-1～44-3)では、パケットバッファメモリ32上の入力インタフェース(Line#1Rx～Line#3Rx)毎のキューから成るキューグループ(42B-11～42B-13、42B-21～23、42B-31～33)において、どの順番でどのタイミングでパケットデータを読み出し、次段のFIFOメモリ45-1～45-3に書き込むかのスケジューリング処理が行われる。

【0067】また、帯域保証クラス用のスケジューラ43-1～43-3では、パケットバッファメモリ32上の入力インタフェース(Line#1Rx～Line#3Rx)毎のキュー(42G-11～42G-13、42G-21～23、42G-31～33)と、ベストエフォートクラス用スケジューラ(44-1～44-3)の次段FIFOメモリ(45-1～45-3)から成るキューグループにおいて、どの順番でどのタイミングでパケットデータを読み出し、次段のパケットバッファメモリ32上のキュー46-1～46-3に書き込むかのスケジューリング処理が行われる。

【0068】キューイングマネージャ33は、帯域保証クラス用スケジューラ43-1～43-3がパケットバッファメモリ32上に送信パケットデータを書き込み、該当する出力インターフェイス用のバッファキュー(46-1～46-3)につなぐと、パケットフォワード31に対してパケットデータの送信処理要求を発行する。

【0069】その他、キューイングマネージャ33は、出力インターフェイス用キュー(送信キュー)46-1～46-3毎にキュー長を監視するキューモニタ48-1～48-3を持っている。

【0070】キューモニタ48-1～48-3と帯域保証クラス用スケジューラ43-1～43-3とはそれぞれ輻輳信号線49-1a～49-3aで結ばれており、キューモニタ48-1～48-3とベストエフォートクラス用スケジューラ44-1～44-3との間は輻輳信号線49-1b～49-3bで結ばれている。

【0071】そして、送信キュー46-1～46-3の長さがあらかじめ設定された閾値を超えると出力インタフェース(Line#1Tx～Line#3Tx)側で輻輳が発生したと判断し、輻輳信号線(49-1a～4

9-3a、49-1b～49-3b)経由で帯域保証クラス用スケジューラ43-1～43-3とベストエフォートクラス用スケジューラ44-1～44-3に対してそれぞれ輻輳通知を行う。

【0072】また、キューの長さがあらかじめ設定された閾値(ヒステリシスを持たせるために、輻輳発生判断に使用する閾値とは違う値を設定する場合がある)を下回ると、通知は解除される。

【0073】ベストエフォートクラス用スケジューラ44-1～44-3と帯域保証クラス用スケジューラ43-1～43-3は、自身に該当する出力インタフェース(Line#1Tx～Line#3Tx)についての輻輳通知を受信すると、スケジューリングアルゴリズムを変更する。輻輳通知が解除されると、通常のスケジューリングアルゴリズムに戻る。なお、変更後のスケジューリングについては、図2及び図3を参照して後に説明する。

【0074】パケットフォワード31は、入力インタフェース(Line#1Rx～Line#3Rx)からパケットデータを受信すると、入力インタフェース番号とパケットの転送先である出力インタフェース番号、通信品質クラス(帯域保証クラス、ベストエフォートクラス)を判断し、パケットバッファメモリ32にパケットデータを書き込んで、該当するキュー(42G-11～42G-13、42B-11～42B-13、42G-21～23、42B-21～42B-23、42G-31～33、42B-31～33)につなぎ、キューイングマネージャ33に対して受信通知を行う。

【0075】また、キューイングマネージャ33からパケットの送信処理要求を受けると、パケットバッファメモリ32上の送信キュー46-1～46-3からパケットデータを読み出し、該当する出力インタフェース(Line#1Tx～Line#3Tx)に対してパケットを出力する。

【0076】図2は、キューイングマネージャ33内部の帯域保証クラス用スケジューラ43-1～43-3の状態遷移図である。

【0077】S10、S11、S21～S2(k+1)、S31～S3(k+1)は、各状態を識別するための記号である[図1の例では、k=3となる]。

【0078】通常はアイドル状態S10にあり、スイッチインタフェースからのパケット受信通知を受けると、ステータスチェック状態S11に遷移する。

【0079】処理対象とするキューグループの内いずれかのキューにパケットデータが存在し、かつキューモニタが輻輳無し状態を示していれば、受信ライン入力インタフェースLine#1Rxのキューからパケットデータを読み出す状態S21に遷移する。

【0080】そこで、パケットデータを先頭からあるバ

書き込み、パケットの最後尾までを読み終わると、次の入力インターフェイスLine#2Rxのキューからパケットデータを読み出す状態S22に遷移し、以降同様に現在の状態に対応する受信ラインのキューからパケットデータを読み出しては次段FIFOに書き込む。

【0081】全ての入力インタフェース数分遷移すると今度はベストエフォートクラス用スケジューラから出力される側のFIFOからパケットデータを読み出す状態S2(k+1)に遷移し、入力インタフェース用キューからの場合と同様にパケットデータを順次読み出しては次段のFIFOメモリに書き込み、パケットの最後尾までを読み終わると、S11の状態に戻る。

【0082】この時点でまだキューグループの内いずれかのキューにパケットデータが存在すると、再度S21に状態遷移する。また、各状態で該当するキューにパケットデータが存在しない場合は、その状態での処理をスキップし、次の状態に遷移する。

【0083】処理対象とするキューグループの内いずれかのキューにパケットデータが存在し、かつ輻輳信号線が輻輳状態を示していれば、入力インターフェイスLine#1Rxのキューからパケットデータを読み出す状態S31に遷移する。

【0084】S31においては、パケットデータを先頭から順次読み出すための読み出しポインタ（パケットバッファメモリ上のアドレスを示す）をある単位例えばバイト単位で移動させる。もし、ポインタがパケットデータの最後尾まで達したならば、該当するパケットデータを全て読み出して次段のFIFOメモリに書き込む。

【0085】あらかじめ設定された回数VAL_L1回までポインタ移動を繰り返すか、またはポインタがパケットデータの最後尾まで達してパケットデータを全て次*

入力インターフェイス#mの

$$\frac{VAL_Lm}{\sum_{i=1}^k VAL_Li} = \frac{\text{トラフィックが占める保証帯域}}{\text{出力インターフェイスの採り得る最大帯域}}$$

(i, k, m はいずれも正の整数)

【0091】

※ ※【数2】

(出力インターフェイスの採り得る最大帯域)

$$\frac{VAL_BE}{\sum_{i=1}^k VAL_Li} = \frac{\text{— (出力インターフェイスの保証帯域)}}{\text{出力インターフェイスの採り得る最大帯域}}$$

(i, k はいずれも正の整数)

【0092】なお、上記式において、「出力インターフェイスの保証帯域」は、各入力インターフェイスにおいて転送される宛先が当該出力インターフェイスのトラフィックの保証帯域の合計と等しい。

【0093】また、状態S11において、処理対象とするキューグループの内いずれかのキューにパケットデー★50

* 段のFIFOメモリに書き込むと、次の入力インターフェイスLine#2Rxのキューからパケットデータを読み出す状態S32に遷移し、以降同様に該当するキューからパケットデータを読み出して次段FIFOに書き込むためのポインタ操作を行う。

【0086】全ての受信インタフェース数分遷移すると今度はベストエフォートクラス用スケジューラ出力FIFOからパケットデータを読み出す状態S3(k+1)に遷移し、受信インタフェース用キューの場合と同様にパケットデータの読み出しポインタ（FIFOメモリ上のアドレスを示す）をある単位で順次移動し、パケットデータの最後尾に達するとパケットデータ全体を読み出して次段のFIFOメモリに書き込み。

【0087】あらかじめ設定された回数VAL_BE回までポインタ操作を繰り返すか、またはポインタがパケットデータの最後尾に達してパケットデータ全体を次段FIFOメモリに書き込むと、S11の状態に戻る。

【0088】この時点でまだキューグループの内いずれかのキューにパケットデータが存在すると、再度S31に状態遷移する。また、各状態で該当するキューにパケットデータが存在しない場合は、その状態での処理をスキップし、次の状態に遷移する。

【0089】VAL_L1～VAL_LKは、該当する各入力インタフェースから入力されるトラフィック（パケット列）が転送先ライン上に占める保証帯域の比率に基づいて決定される。また、VAL_BEは、転送先の出力インターフェイスの取り得る最大帯域と合計保証帯域の差分に基づいて決定される。VAL_L1～VAL_LK, VAL_BEは以下の関係が成り立つ。

【0090】

【数1】

★タが存在しない場合には、アイドル状態S10に遷移する。

【0094】図3は、キューイングマネージャ33内部のベストエフォートクラス用スケジューラ44-1～44-3の状態遷移図である。図2に示す帯域保証クラス用スケジューラが取り得る状態から状態S2(k+1)

17

および状態S3(k+1)を削除した内容となっている。また、出力インターフェース輻輳状態でのキュー読み出し回数が異なっている。

【0095】VAL_BE_L1~VAL_BE_LKは、該当する各入力インターフェースから入力されるトラフィック（パケット列）が入力インターフェイス上で占*

$$\frac{\text{VAL_BE_Lm}}{\sum_{i=1}^k \text{VAL_BE_Li}} = \frac{\text{（入力インターフェイス\#mの採り得る最大帯域）} - \text{（入力インターフェイス\#mの保証帯域）}}{\text{全ての入力インターフェイスに渡る分子の合計}}$$

(i, k, m はいずれも正の整数)

18

*める保証帯域の比率に基づいて決定される。VAL_BE_L1~VAL_BE_LKは以下の関係が成り立つ。

【0096】

【数3】

【0097】なお、ベストエフォートクラス用のスケジューラ(44-1~44-3)及び帯域保証クラス用のスケジューラ43-1~43-3は、それぞれ対応するキュー(42G-11~42G-13、42B-11~42B-13、42G-21~23、42B-21~42B-23、42G-31~33、42B-31~33)から上述の説明にしたがったアルゴリズムでパケットデータを読み出すが、読み取ったパケットデータがパケットの途中である場合には、次段のFIFOメモリ(45-1~45-3)或いはキュー(46-1~46-3)に読み取ったパケットデータを出力せず、完全にパケット全体を読み取った後に、次段のFIFOメモリ(45-1~45-3)或いはキュー(46-1~46-3)にパケットデータを出力する。

【0098】したがって、本実施の形態のネットワーク間中継装置によれば、出力インターフェースの輻輳が検出された場合に、輻輳が検出された出力インターフェースに対応する各スケジューラがパケットの取り出しレートを変更することができるので、入力インターフェース毎に通信品質が要求されるデータの最小帯域を保証することができる。

【0099】図4は、パケットバッファメモリ上のパケットバッファキューの構造を示す。

【0100】キューはパケットデータが格納されるj個のバッファとその先頭ポインタ（メモリ上のアドレス）を示すキューディスクリプタから成る。

【0101】キューディスクリプタには、各パケットバッファに対応したj個のエントリが存在し、各エントリはバッファ先頭ポインタを示すフィールドとバッファ内のパケットデータサイズを示すフィールドから成る。これらのエントリはサイクリックに使用される。

【0102】そのほかキューディスクリプタはキューイングされたパケットデータの合計であるトータルデータサイズを示すフィールドを持ち、またキューイングマネージャ内のスケジューラとDMAコントローラがキューを操作する際に両者の間でディスクリプタ・アクセスの排他を行うためのセマフォ・フィールドを持つ。

【0103】キューイングマネージャ内には、キューデ※50

※ィスクリプタの先頭ポインタと末尾ポインタを格納するレジスタが存在し、また実際のキューの先頭エントリと末尾エントリのそれぞれの先頭ポインタを格納するレジスタが存在する。

【0104】このキューの先頭エントリと末尾エントリのそれぞれの先頭ポインタが示す位置の間に存在するエントリがキューに相当する。また、先頭エントリと末尾エントリのそれぞれの先頭ポインタ値が等しい場合には、キューイングマネージャはキューにパケットデータが存在しないと判断する。

【0105】各種スケジューラがパケットデータを読み出す際や、パケットフォワードへパケットデータを出力するためにDMAコントローラがパケットデータを読み出す際には、先頭エントリのポインタ格納レジスタを参照してキューディスクリプタ上の先頭エントリの位置を割り出し、続いてエントリ内のバッファ先頭ポインタからキューの先頭パケットデータの位置を導き出す。

【0106】パケットデータの読み出し終了後、先頭エントリのポインタ格納レジスタの値を次のエントリのポインタの値に置き換える。

【0107】パケットフォワードからパケットデータを入力するためにDMAコントローラがパケットデータを書き込む際には、末尾エントリのポインタ格納レジスタを参照してキューディスクリプタ上の末尾エントリの位置を割り出す。キューの末尾に続く次のエントリがフリー・バッファの位置を示すため、そのバッファにパケットデータを書き込む。

【0108】パケットデータの書き込み終了後、末尾エントリのポインタ格納レジスタの値を次のエントリのポインタの値に置き換える。

【0109】また、キューディスクリプタ上のトータルデータサイズ・フィールドにパケットデータサイズ分を加算する。

【0110】逆に、各種スケジューラがパケットデータを読み出す際や、パケットフォワードへパケットデータを出力するためにDMAコントローラがパケットデータを読み出す際には、トータルデータサイズ・フィールドにパケットデータサイズ分を減算する。

【0111】このトータルデータサイズ・フィールドの値が0でなければ、当該キューにパケットデータが存在することになる。

【0112】すなわち、本実施の形態においては、各スケジューラは、入力キューからのデータの読み出しの際に、パケットデータを読み出すポインタをある単位で移動させ、ポインタがパケットデータの最後尾に達するまではキューから出力しないで、1度に読み出す読み出し単位をの数を減らすことで、読み出しレートを制御する。

【0113】なお、本実施の形態においては、帯域保証クラス用のスケジューラ及びベストエフォートクラス用のスケジューラによってスケジューリングをする場合について説明したが、ベストエフォートクラス用のスケジューラによる取り出しレートの変更は必ずしも必須の構成要素ではない。

【0114】すなわち、ベストエフォートクラス用キュー(42B-11~42B-13、42B-21~23、42B-31~33)にパケットデータが格納されると、ベストエフォートクラス用のスケジューラによって取り出しレートの変更をすることなく、各キュー(42B-11~42B-13、42B-21~23、42B-31~33)から順番にパケットデータが取り出されて、対応するFIFO(45-1~45-3)に格納される構成としてもよい。

【0115】この場合、取り出しレートの変更は帯域保証クラス用のスケジューラのみで行なわれることになる。

【0116】したがって、本実施の形態のネットワーク間中継装置によれば、出力インタフェースの輻輳時でも、入力インタフェース毎に設定された帯域を保証し、かつ当該出力インタフェースの残りの帯域を入力インタフェース毎のベストエフォートトラフィックの割合で割り当てることができる。

【0117】結果的に、入力インタフェース毎に最小帯域の保証が可能となり、かつ入力インタフェースの限界帯域に応じた割合でパケット中継を行うことにより、公平さを保つことができる。

【0118】また、可変長パケットを転送するネットワークにおいて、中継装置の入出力インタフェース間の最小帯域を保証することにより、その区間を経由する通信の最低限のスループットを予測することができるようになる。

【0119】なお、本願発明は、上記各実施形態に限定されるものでなく、実施段階ではその要旨を逸脱しない範囲で種々に変形することが可能である。また、実施形態に示される全構成要件から幾つかの構成要件が省略されることで発明が抽出された場合には、その抽出された発明を実施する場合には省略部分が周知慣用技術で適宜補われるものである。

【0120】また、実施形態に記載した手法は、計算機(コンピュータ)に実行させることができるプログラム(ソフトウェア手段)として、例えば磁気ディスク(フロッピー(登録商標)ディスク、ハードディスク等)、光ディスク(CD-ROM、DVD、MO等)、半導体メモリ(ROM、RAM、フラッシュメモリ等)等の記録媒体に格納し、また通信媒体により伝送して頒布することもできる。なお、媒体側に格納されるプログラムには、計算機に実行させるソフトウェア手段(実行プログラムのみならずテーブルやデータ構造も含む)を計算機内に構成させる設定プログラムをも含む。なお、本明細書でいう記録媒体は、頒布用に限らず、計算機内部あるいはネットワークを介して接続される機器に設けられた磁気ディスクや半導体メモリ等の記憶媒体を含むものである。

【0121】

【発明の効果】以上詳記したように、本発明によれば、可変長パケットを中継するネットワーク間中継装置において、重要度の高いデータの最小帯域を保証することができるネットワーク間中継装置を提供することができる。

【図面の簡単な説明】

【図1】本発明の実施の形態に係るネットワーク間中継装置のパケット処理系を示す図。

【図2】キューイングマネージャ33内部の帯域保証クラス用スケジューラ43-1~43-3の状態遷移図。

【図3】キューイングマネージャ33内部のベストエフォートクラス用スケジューラ44-1~44-3の状態遷移図。

【図4】パケットバッファメモリ上のパケットバッファキューの構造を示す図。

【図5】ネットワーク間中継装置であるルータの一般的な内部ブロック図。

【図6】従来のネットワーク間中継装置のパケット処理系を示す図。

【図7】スケジューラ4-1~4-3の動作を説明するための状態遷移図。

【符号の説明】

3-11~3-13、3-21~3-23、3-31~3-33...キュー、4-1~4-3...スケジューラ、5-1~5-3...キュー、11-1~11-3...PMD/PHY、12-1~12-3...データリンクコントローラ、13...パケットフォワーダ、14...パケットバッファメモリ、15...キューイングマネージャ、31...パケットフォワーダ、32...パケットバッファメモリ、33...キューイングマネージャ、41-1~41-3...FIFO、

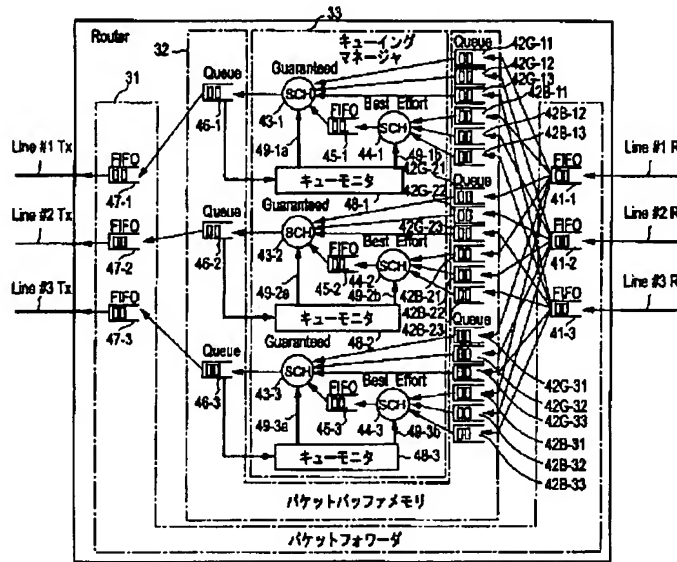
21

42G-11~13、42G-21~23、42G-3
1~33...帯域保証クラス用のキュー、
42B-11~13、42B-21~23、42B-3
1~33...ベストエフォートクラス用のキュー、
43-1~4...帯域保証クラス用のスケジューラ、
44-1~4...ベストエフォートクラス用のスケジューラ、

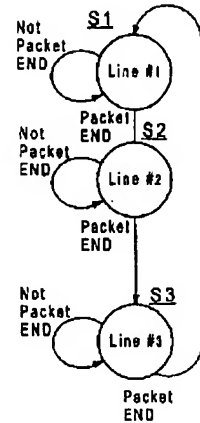
22

45-1~3...FIFOメモリ、
46-1~3...キュー、
47-1~3...FIFO、
48-1~3...キューモニター、
49-1a, 1b, 49-2a, 2b, 49-3a, 3
b...転換信号線。

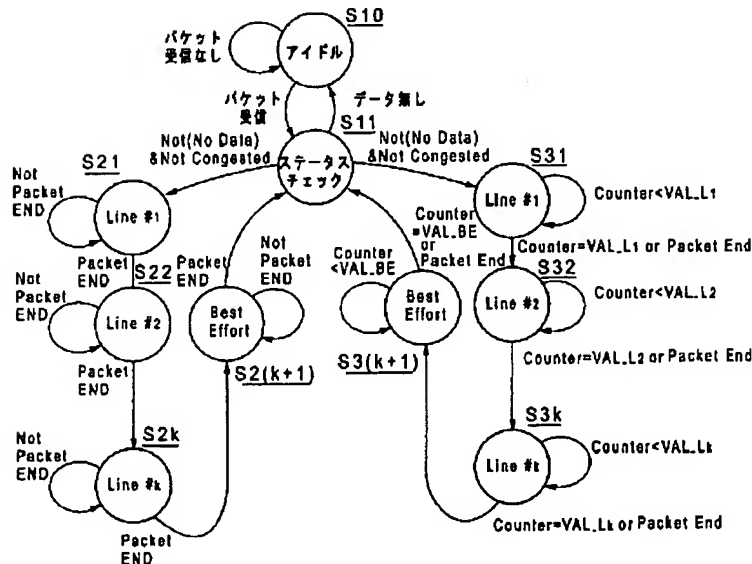
【図1】



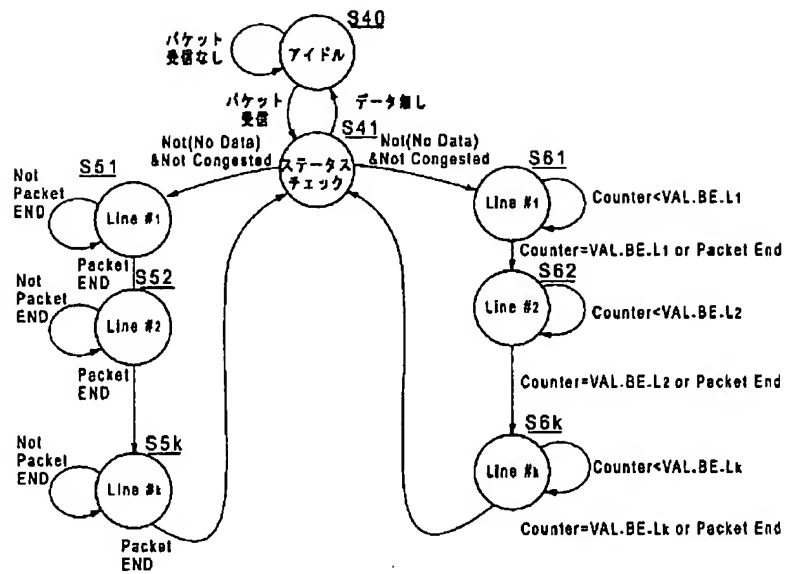
【図7】



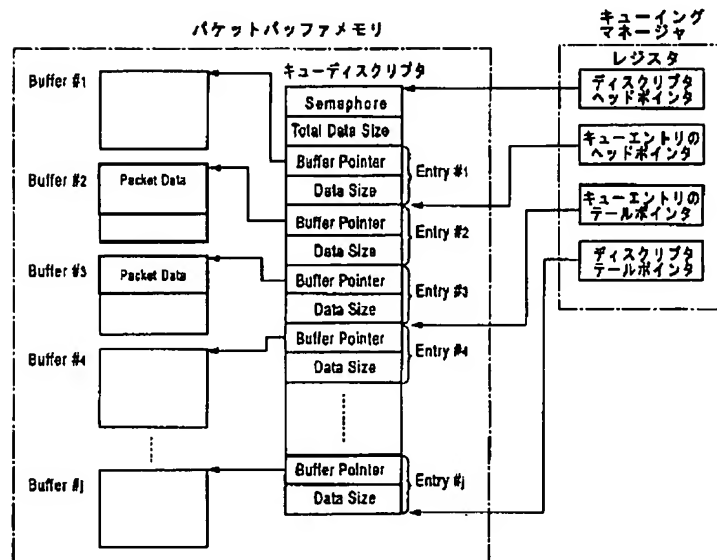
【図2】



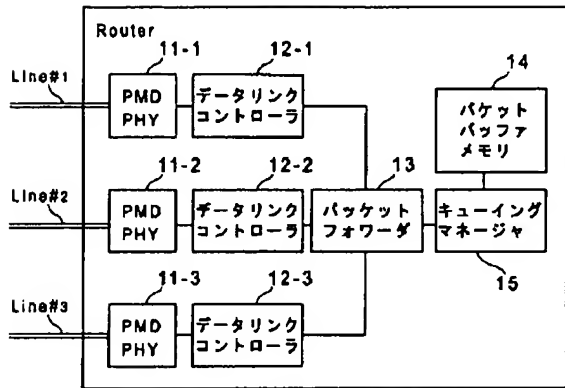
【図3】



【図4】



【図5】



【図6】

